

THE CONVENTIONAL WISDOM OF BEHAVIOR ANALYSIS: RESPONSE TO COMMENTS

J. E. R. STADDON

DUKE UNIVERSITY

The willingness of behavior analysts to entertain criticism from any quarter is one of the best kept secrets in psychology. This openness stands in sharp contrast to the confident pronouncements of many thoughtless critics. Try, for example, the following from the frontispiece of the dominant textbook in neuroscience. The color picture shows a PET scan of a human brain engaged, we are assured, in thought. The caption opines thus:

Behaviorism dominated experimental psychology for a good part of the 20th Century. Behaviorists thought [we should, I suppose, be grateful for the verb if not for the past tense!] that the only way to study behavior was by examining a subject's observable actions. They regarded the brain as an unapproachable black box and denied the usefulness of studying mental processes because they were basically unobservable. The current view of psychology is very different. Most psychologists now want to look into the black box and understand how mental processes function. (Kandel, Schwartz, & Jessel, 1991, p. 3)

The paragraph is of course a small masterpiece of misrepresentation and non sequitur. (Can we really study behavior other than by observing behavior? Are pictures of brain activity the same as pictures of mental activity? Do we have evidence of "mental activity" other than via "observable actions"? Did Clark Hull, a behaviorist, regard the brain as an "unapproachable black box"? And so on . . .) In the authors' defense, this is a textbook, after all: Philosophical subtlety is not required, and brain scans are great propaganda. We cannot expect neurobiologists to be entirely *au courant* with what is happening in psychology, and one of the collaborators in the PET picture is a cognitive psychologist. Nevertheless, the fact

remains that behaviorism is viewed as irrelevant to the cognitive and neuroscience "revolutions" by most psychologists as well as by those neurobiologists who think much about behavior.

This view is wrong. Most behaviorists are happy, even anxious, to include brain function in their conceptual schemes (see, e.g., John Donahoe's comments here). Even B. F. Skinner, at least in the early days, averred that explanatory "concepts which do not make physiological formulation impossible and which are amenable to growing physiological knowledge are preferable, other things being equal, to those not so amenable" (1938, p. 440). Many behavior analysts study aspects of cognitive function (see Charles Shimp's comments), even if they do not use the "cognitive" label. How, then, have behaviorists—so open to criticism, so willing to entertain alternative views—been so effectively painted into a philosophical corner?

I think that part of the answer can be found in some of the reactions to my critique. Charles Catania's response is the most direct. He begins his commentary, "In a nutshell, my problem with Staddon's essay is that it endorses an implicit philosophy of science that is not behavior analytic" (p. 449). He may be right, but this is nevertheless a curious objection. To see why, imagine a critique of Freudian theory that is met with the response "Eysenck's implicit philosophy of science is not psychoanalytic." How persuasive would such a response be to a non-Freudian? How open to argument and evidence would such a respondent appear to a neutral observer? The problem with an objection like this is obviously that it begs the question of whether or not the philosophy under dispute is *correct*. Indeed, it implies that a non-behavior-analytic philosophy cannot be correct. An uncharitable cognitivist (there are a few) might well interpret it as an assertion of the *inerrancy* of this particular view of science and of its philosophy—a stance usually associated not with science but with religion.

I thank members of the Learning and Adaptive Behavior Group for many vigorous discussions of these issues over the years. Research supported by NSF and NIMH. Address correspondence to J. E. R. Staddon, Department of Psychology: Experimental, Duke University, Durham, North Carolina 27708-0086.

Behavior-Analytic Philosophy

Nevertheless, we can ask: What is behavior-analytic philosophy? How is it different from Staddon's philosophy? Is it correct? Catania, following Skinner, explains that "A behavior-analytic philosophy of science must begin not with assumptions about truth and knowledge but rather with the behavior of the scientist" (p. 449). This affirmation has two parts: Behavior analysis encompasses both the behavior of the subject and the behavior of the experimenter, and it "must begin" with the behavior of the experimenter. Catania's specific objection is that my philosophy of science does not say how terms like *explanation*, *theory*, *model*, and *fundamental knowledge* "bear on the behavior analyst's behavior when the behavior analyst explains or theorizes or models or knows" (p. 449). This is a plausible objection, because the experimenter is obviously an organism subject to the same behavioral laws as other organisms. Surely the burden of proof must be on those who reject the idea that the laws of behavior apply to the experimenters and their subjects alike?

There are nevertheless three problems with this view of behavior analysis. The first is that it repudiates an assumption that is absolutely intrinsic to the rest of natural science, namely the separation of subject and object: the clear separation of the scientist from the object of his investigations. Catania is not alone in wishing to abandon the subject-object distinction: Steven Hayes's discussion of contextualism and Timothy Hackenberg's concluding paragraph point in exactly the same direction. The second problem is that Catania's view assumes that we already know enough about the laws of behavior—derived largely from behavior of pigeons and rats—to apply these laws *now* to ourselves as scientists. The third problem is that the refusal to separate subject from object ignores the practical difficulties entailed by considering both halves of the "epistemological equation" at the same time. Let's look at these objections one by one.

The assumption that observer and observed can be considered separately, that in the study of organic chemistry we do not need at the same time to consider the fact that we are ourselves organic compounds, underlies all of natural science. Yet we are told that behaviorists, alone in science, should give it up. What

compels this conclusion? Has the experimental analysis of behavior been such a stunning scientific success that we can confidently overturn an axiom that underpins the rest of science? What would a disinterested observer say to such a proposition? I don't think many outside the experimental analysis of behavior—and perhaps few within—would find it very persuasive.

The second problem is our present level of understanding. How much do we really know—scientifically, not as species members—about the behavior of human beings? Do we understand why some children grow up as geniuses and others as dullards? Do we know what makes a hero or a coward? Do we know why people believe in religion: Are the essentials really captured by an experiment in which a hungry pigeon bobs and pecks on a free-food schedule? Is "scientific behavior" different from "nonscientific behavior," and, if not, why did science evolve in Europe and not in China? Is there anything special about human language, and, if so, what is it? Is there such a thing as morality? Was Woody Allen's behavior simply "inappropriate," as his shrink called it, or was it wrong? Why do all cultures *have* concepts of right and wrong? What is the scientific status of these concepts, or are they perhaps not scientific at all?

These are ambitious questions, perhaps. (I attempt to address some of them in a forthcoming book on behaviorism). They are nevertheless invited by any claim that our current science of behavior not only can but must be applied *now* to one of the most complex activities of which human beings are capable, namely the advancement of scientific knowledge. But simpler questions also suffice to show the limitations on our current understanding: Why is punishment sometimes effective and improving and sometimes ineffective and degrading? Likewise for positive reinforcement: What makes something reinforcing? Why do some people like Coke® and others Pepsi®? Is "reinforcement history" really a sufficient explanation for every nook and cranny of human behavior? How does it explain individual differences, for example (what about Mozart)? Armando Machado describes a number of other research areas that behavior analysts have largely ignored. At a nuts-and-bolts level, researchers are still arguing, more than 30 years after its discovery, about the causes of behav-

ioral contrast, a supposedly basic phenomenon. Papers on the matching law, a discovery of similar vintage, continue to appear in the pages of this journal. How far have we come, really? I submit that our present understanding, far from being something to be "given away" to a benighted public, is fragmentary and often very misleading. (I include most of psychology in this; behavior analysis is not worse than many other areas of psychology—indeed, it is better than most.) To apply it to ourselves as experimenters when we are far from sure of its scope even for rats and pigeons, strikes me as highly premature—and dangerous, because it assumes a level of understanding that does not in fact exist. By assuming behavior analysis to be not only true but complete, we close our eyes to alternative views and to gaps that seem obvious to outsiders. By all means apply a science of behavior to the behavior of scientists, but please, let's wait until we have a science rather than the beginnings of one.

The final problem is simply a practical one. The natural-science separation of subject and object may have worked so well simply because it breaks down problems into manageable parts. We have no evidence from any other science that refusing to separate subject from object can work as well. This is a pragmatic defense that may appeal to behavior analysts when all others fail.

Truth and Explanation

The version of behavior analysis described by Catania, Hayes, and Hackenberg seems to represent a radical break with practice in other sciences. The argument is perhaps clearest in Hayes's commentary, when he writes "truth is a matter of successful working . . . [which] means the production of a specified consequence" and "prediction and control are the consequences . . . important for [Skinner's] science" (p. 461). The focus is on action and environmental change. This view seems to rule out sciences like astrophysics and paleontology, where the subject matter cannot be manipulated. But even a novice behavior analyst would surely respond that in these sciences, the "consequence to be produced" is just "successful prediction." "Successful working" for an astrophysicist then becomes simply making successful predictions.

Yet this escape is too successful, because it erases any real difference between the behavior

analyst's view of science and everybody else's, and raises the question: If good astrophysics is the generation of successful predictions, why should good psychology not proceed in the same way? Once again, our resourceful student might respond, "The difference is that behavior analysis is an *experimental* science; astrophysics is not. In an experimental science, control takes precedence over prediction." According to Hayes, behavior analysts believe, with Skinner, that scientific knowledge is a corpus of rules for effective action. This is simply the pragmatic view of truth, about which philosophers have written at length (see also Stadon, in press, chap. 3 and 4).

There are a number of difficulties with plain-vanilla pragmatic epistemology, but the one I will mention here is its lack of a time horizon: "successful working," yes, but when? A real difference between science and craft is the time the practitioner is willing to wait before achieving power over nature. A Renaissance painter, anxious to improve the quality or durability of his paints, would obviously have done better to operate by trial and error and the rules of thumb of the painter's trade than to wait on the development of scientific chemistry. Indeed, a strict application of Hayes's "successful working" rule might well have prevented chemistry from developing at all. Successful working *now* may be the enemy of much more successful working down the road. The lesson I take from the history of science is that a search for *understanding* (no matter how ill-defined that idea may be) will eventually yield much more control over nature than a shortsighted emphasis on control now. American businesses are often criticized, and contrasted with their counterparts in Japan, for their overemphasis on the next quarter's profits. It would be unfortunate if American behavioral science were to fall victim to the same shortsightedness.

The pragmatic emphasis of some behavior analysts at times seems to imply a renunciation of "truth" as a scientific ideal: "Analysis ends not with a discovery of the truth, but with the production of verbal constructions that help achieve an effect" (Hayes, p. 461). Skinner's occasional dismissals of logic can be read in the same way, and recent writings in nonscience areas such as critical literary theory are perfectly explicit in their denial of objectivity. There are two obvious problems with the

abandonment of truth, one ideological and the other practical. The ideological objection is because "the idea that there is no such thing as objective truth . . . [is] closely linked with authoritarian and totalitarian ideas" (Popper, 1962, pp. 4–5). The practical objection is that belief in a single "real world" independent of ourselves seems to have been essential to the historical development of natural science. Because a belief in "truth" entails no more than this, we abandon it at our peril, it seems to me.

Internal State

Several commentators question my concept of *internal state*. Some of these objections seem to reflect a misunderstanding of what I was trying to say. Others proceed from different meanings given to the words *internal* and *theory*. Still others reflect real philosophical differences.

First, definitions: Hard-core behavior analysts (you know who you are!) accept Skinner's conclusion that there is nothing special about the skin as a boundary—"private" events and public events are to be treated in the same way. I cannot agree, because I cannot understand what Skinner means (for example) when he speaks of imagining a piece of music as "a response." What (or who) is responding? I'm afraid that I will always feel that the skin *does* make a difference: *behavior* is what we, third-party observers, can see from outside. Whatever happens inside is either private (hence, not directly accessible to anyone else) or internal (some property of the nervous system).

I have tried to make some sense of all this in *Behaviorism* (Staddon, in press), where I propose, as a matter of definitional discipline, that we divide the world into three domains: Domain 1 is just the phenomenological world of felt, private experience. Philosophers call private sensation *quale*, and we notoriously know only our own *qualia*. This is the truly private world. Domain 2 is the public world of physiology: our intersubjectively verifiable, shared view of the rat's brain. Domain 3 is the public world of behavioral (i.e., outside-the-skin) data. Domain 3, which includes verbal reports as well as lever presses, provides the raw data for behavioral psychology.

Internal states fit into this scheme in the following way. They are not Domain 3 be-

havior, because they cannot be measured directly. They are not Domain 2 behavior, because they do not involve brain measurements. And they are certainly not Domain 1, because they are not private and not mental. They are thus not subject to Skinner's objections to internal states, which he conceived of as either mentalistic or physiological. So what do I mean by the term? Internal states in my sense are not behavior but *inferences* from behavior (perhaps we should call them *infernal states*!). As I have argued elsewhere (Staddon & Bueno, 1991) their relation to Domain 2, the domain of nervous system structure and function, is more or less uncertain depending on whether the theory that embodies them is fragmentary and tentative or comprehensive and well established: The better the theory, the closer the relationship to physiology is likely to be. There is nothing especially puzzling about this uncertainty. It is shared by every theoretical concept in its early days. The units of inheritance implied by Mendel's empirical laws were conjectural for 50 years; the physiological reality of Sherrington's synapse took a comparable time to be confirmed. Quarks have passed from conjecture to reality in the last two decades. "A word that begins as a theoretical term can evolve into a name for a phenomenon," in Catania's words (p. 451).

Thus, Hackenberg's attribution to me of the "implicit assumption, dating back to Socrates, that the formal structure of a theory corresponds with the formal structure of the world" (p. 459) is not correct. Contra Donahoe and Carol Pilgrim, I do not assume that the internal states of the CE model (say) correspond to the world (i.e., Domain 2 properties of the nervous system), although I hope that they do. As I point out in my essay, a theory may be useful for several reasons even if it turns out to be false (as most do), so the correspondence or lack of correspondence between the states of a model and states of the nervous system is not in any way crucial to the model's utility. Nevertheless, I am confident that if someone does come up with a truly predictive and comprehensive theory of operant behavior, some aspects of that theory will have nervous-system counterparts. This is to say no more than that the behavioral function of the nervous system must bear some relationship to its physiological and anatomical properties.

Miscellaneous

These three topics—behavior-analytic philosophy, truth and explanation, and the idea of internal state—seem to be the main points of contention. I turn now to the more specific points.

Theory. Catania seems to use the word *theory* as an antonym for *fact*, as in “At what point are correspondences [between data and theory] close enough that a model is no longer theoretical?” (p. 451). This may just be a matter of nonstandard word usage. Major theories like quantum theory, relativity, or the theory of evolution are as certain as the facts on which they rest. In some respects, they are more certain, or at least more enduring. No one remembers what balls were rolled down what inclined planes to provide data on which Galileo and Newton could base their theories. Yet Newtonian physics, though false, strictly speaking, is still taught to beginning physics students. Contra Skinner, facts are no more enduring than theories. In a too-famous comment, he wrote, “Most theories are eventually overthrown, and the greater part of the associated research is discarded” (1950, p. 270). The truth is that almost everything in science is discarded, facts as well as theories. Skinner could with equal justice have written, “Most facts are eventually subsumed under theories, and are then discarded.”

Catania also argues, in good Machian fashion, that mathematical theories or models should be treated not as explanations but as “economical descriptions.” This seems to me a matter of labeling. If you want to consider quanta as “economical descriptions” rather than real entities, fine. “Reality” is only as real as one’s epistemology permits. If someone wishes to consider the three-dimensional world in front of his eyes no more than an “economical description” of a succession of two-dimensional binocular images, we may consider him eccentric, but we cannot prove him wrong. The point is not whether a theory is a description or an explanation but whether it can withstand the usual scientific tests.

I am puzzled by Catania’s statement that “a behavioral philosophy of science should treat theories not as causes of scientific behavior but as its products” (p. 451). I am puzzled because in one sense he is obviously correct: Of course

theories are “products” of “scientific behavior”; of what else could they be products? But on the other hand, once produced, they also guide scientific behavior (check out the quote from Skinner, above, who obviously felt that theories not so much guide as *misguide*). I don’t suppose that Catania would contest that either. Perhaps we don’t differ on this one.

Initial conditions. Hayes criticizes the CE model because “it does not say precisely where the initial values come from” (p. 463). I don’t understand this objection. Any “black box” has an initial state, whose details are not normally written on the outside. All that can be required of any theory is that the theory specify what needs to be done—what stimuli presented, what measurements made—to infer the initial conditions. For some deterministic systems, the initial conditions are literally unobservable: No measurements at times t_1 through t_N will suffice to identify unambiguously the system state at time t_0 (the beginning of the experiment), even if an indefinite number of system replicas are available. For the system represented by the CE model, there are, in fact, simple tests that will allow one to estimate the four state variables at t_0 . What more is required?

Was Skinner right after all? Richard Shull asks about the relation of my idea of state variable to state variables like *response strength* proposed by Skinner. There are some similarities, although even concepts like the *reflex reserve* were treated by Skinner as static entities, rather than as components of a dynamic process model. The CE model does have a concept like response strength, the V values of the competing responses. But V has two dimensions, cumulated responses as well as cumulated reinforcers, rather than just one. On the other hand, responses compete strictly according to the single V value, so that V values share at least this property with response strength. A big difference, though, is that response rates for two concurrently available responses may differ widely, even though their V values are almost identical (this is a consequence of the winner-take-all response rule and the fact that V has two dimensions).

But the main difference between Skinner’s early views and what I propose is that Derick Davis and I have suggested an actual process by which response strength is altered and by which different responses compete. Skinner al-

ways drew back from constructing an actual learning process, a real model. I think this is because his lifelong fundamental purpose was not scientific but *meliorative*. He wanted to improve mankind. He may have been interested in the study of behavior "in its own right"; he certainly was not interested in pigeons and rats "in their own right." His aim was better living through behavior analysis. He may have judged that he could achieve it more swiftly through experimental analysis than by theoretical exploration. And given the explosion of new phenomena revealed by the "Skinner-box" technology and the primitive means then available for theoretical exploration, who can say that he was wrong?

But today the balance has shifted. We have more data, and truly novel results now rarely emerge from purely inductive experimentation. Computers permit the easy exploration of processes whose properties could hardly be grasped in a lifetime of study by gifted mathematicians 40 years ago. It is time for theoretical behaviorism to explore some avenues Skinner himself glimpsed, but did not enter, many years ago.

In conclusion, I thank all the commentators for taking time to reexamine some very old

issues. I apologize to those whose criticisms space prevents me from addressing, and I congratulate those whose perspicacity led them to agree with me. I thank Jack Marr and Rick Shull for educating me about Mach, Armando Machado for correcting me about Darwin, and Peter Killeen for reminding me that definitions need some stability. Above all, I thank Marc Branch for the *JEAB* editorial that provoked my initial response and Phil Hineline for his patience and good will in orchestrating the whole thing.

REFERENCES

- Kandel, E. R., Schwartz, J. H., & Jessel, T. M. (Eds.). (1991). *Principles of neural science* (3rd ed.). New York: Elsevier.
- Popper, K. (1962). *Conjectures and refutations: The growth of scientific knowledge*. New York: Basic Books.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, **52**, 270-277.
- Staddon, J. E. R. (in press). *Behaviorism: Mind, mechanism and society*. London: Duckworth.
- Staddon, J. E. R., & Bueno, J. L. O. (1991). On models, behaviorism and the neural basis of learning. *Psychological Science*, **2**, 3-11.